



RECEIVED
FEB 14 2002
148
FBI CENTER 160012900

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/727,855B

DATE: 02/06/2002
TIME: 09:42:42

Input Set : A:\PTO_VSK.txt
Output Set: N:\CRF3\02062002\I727855B.raw

PS

3 110> APPLICANT: HOSHINO, Tatsuo
4 OJIMA, Kazuyuki
5 SETOGUCHI, Yutaka
7 120> TITLE OF INVENTION: PROCESS FOR THE MANUFACTURE OF CAROTENOIDS AND BIOLOGICALLY
USEFUL

8 MATERIALS THEREOF
10 130> FILE REFERENCE: C38435/111694
12 140> CURRENT APPLICATION NUMBER: 09/727,855B
13 141> CURRENT FILING DATE: 2000-12-01
15 160> NUMBER OF SEQ ID NOS: 17
17 170> SOFTWARE: PatentIn version 3.1
19 210> SEQ ID NO: 1
20 211> LENGTH: 3632
21 212> TYPE: DNA
22 213> ORGANISM: Phaffia rhodozyma
24 400> SEQUENCE: 1

ENTERED

25	tcctgttgat	aatctttcta	acgccttgta	ctttgaccaa	ggcgtttgtc	cgaaattttg	60
27	caaaacttagt	gttggtcgca	tggacggtct	tgggatccag	aactgacggc	tggccaataa	120
29	agtatgacga	tggtagaggt	gaaggaggga	accacaggtt	gaccagtctc	aaagagtgtc	180
31	gatgtgcgcg	aggatttgta	attaaatggt	gttgtatatg	ctagagccaa	gagaagacat	240
33	ttgggttttg	ttttggtttt	gcatttgatg	agatgtgtca	cgattgaaga	cgggaggagg	300
35	ctcactaacc	caagaagcca	ggatcaggag	gaatgcctcc	cccttttcat	caagatcttt	360
37	ctcacatcga	acatttgaca	ttctctttag	tatctttcta	tccttttctt	ccaactttct	420
39	ccattgtatc	gactttgtct	gacttgcctc	tcttatctct	gagcagagat	gggcattcca	480
41	atategaagg	agcgacacaa	gaccttgagg	tttgggtaac	agatgaagag	gggcagagggt	540
43	ggatggggct	gtaggaagta	gctgatcgat	gagttcctgg	atgatgatag	gcgaaggaa	600
45	agacatagga	tctctgtctc	gtcctggaat	tactgagtct	tgtatccagc	gtgttcttgt	660
47	ctcgaagaag	ccttcaagat	cgatgtaaga	taagacaggc	aatgaggagc	gacgaacgaa	720
49	cgaacgaaaa	gaacagaaga	gctggtaagt	cagtcagtca	gtcagtcagt	caatcaaaac	780
51	ctgggtgtcta	gggtttatagc	tcgacgcgac	gcgacgcgtt	tgagacgcga	tatgcttaag	840
53	taatacctgg	cgtcatcccc	ccagccgagg	caagagccga	gccgctcgtg	aacgacaaaa	900
55	ttcaaaaaggc	tttctccatc	ttaaagctcat	tctcatctaa	ccgactcctc	tcgttcccat	960
57	cattcccate	attctacgcg	catccatgtc	tgttcgagca	tcctctctct	ccgtgtctag	1020
59	acagactttc	gtcgtcctcg	ctgctttcca	gatcagggca	aagcataccc	tgccgtgagct	1080
61	tccttaagct	taogatgtaa	gacttttccg	tgttctctca	ttcgtcgtct	tcttggtttt	1140
63	tttctgtctc	gcctcttagc	tcttcttctg	cctttctgtc	ctgctctttg	ttgttgatat	1200
65	tcagctcgat	agaactaacc	atctcatctc	ctggacattc	ttttaactgga	aacgtatctt	1260
67	gtccttggtt	tttcttggtt	ttggttgaaa	attctctctc	actcaggccc	tggagccctc	1320
69	catctccaag	gagatcatga	cccttcacca	caccaagcac	catcagaact	atgttaacgg	1380
71	cctcaacgct	gccgaggaga	gctaactcgc	cgtctgtggc	aaggaggatg	tgcttaacca	1440
73	ggttaagctt	cagtctgtac	gtctgacgct	ttttttatcg	accggaacgc	ctggtgagga	1500
75	gggagatgaa	gtttgatgag	cgtctcatct	ctagcaggtt	gacccgatca	tacaggctct	1560
77	caagttcaac	ggaggaggac	acatcaatcg	tcagtgatat	tcttcaaaact	cttgctgagc	1620
79	aagtcaggtc	aagctgactg	tttctgcttg	tttctgcgga	tctatctcat	ccttgatttg	1680

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/727,855B

DATE: 02/06/2002
TIME: 09:42:42

Input Set : A:\PTO_VSK.txt
Output Set: N:\CRF3\02062002\I727855B.raw

```

81 gcatgatgaa acagactctc tgtttctggaa gaacttggct cccataggat ccgaggaggg 1740
83 taccctctct gaaggacctc tcaagaaggc tctcgaggaa tcttttgggt ctttcgaggg 1800
85 tccgtccatc tatcttctca ttcagttgtg tttggttccg gtatactcat ctgttttgtt 1860
87 tccccacaaa ataaaaataa aaatcttgtc ctctccgggg gtccgactgc acgttcatag 1920
89 ccttcaagaa gaagttcaac gctgacaccc ctgctgtcca aggatccgga tggggctggc 1980
91 ttgtatgtat catatccttt ccatctcaaa ctcttctcag agtctttttc cttgagactt 2040
93 caaactgaat atacatgttt ctacaacaaa caacaggggc tgaaccgct tactaagaag 2100
95 ctggaagtca ccacgaccgc caaccaggac cctctgctta gtaagtgtt tctacatgat 2160
97 tttctatctc aacgcgatct gcatgattcg tcactgattc actggattct cttgtttcgt 2220
99 tttctcggg atgatttcat aaacagctca cttctctatc atcggagtgt acgtgcgtat 2280
101 ctttcttgaa tagtctgagc gtctgatctc gttttattga ctgacgtgtt gcttctgtcc 2340
103 aaatcattaa aaaaaatgaa aacaaataat cgattgaccg acgaaaacag atctgggagc 2400
105 agcttttcta ccttcagtac aagaacgtca agcctgacta tctcgtctgt gtttggtccg 2460
107 ttatcaacta caaggaggca gaggcctgat tgcaggctgc tctctaagcg ggacgaaaaa 2520
109 taacgacata tgaagggagg atcaaatatc gtttcttcat aaacaacttt cgaggcagat 2580
111 gggagagtac gtacaagaga ggtttgtatg gagaattgag tttgttgacg gtttagcaggt 2640
113 tatgatatat gtagctatag tctagtctaa atctgaaaga agagaacaag atggtttgtc 2700
115 cgaagagatt gagagatcaa gcccggtcat ctgatgtcga acaaacatgc cctggtctgc 2760
117 caacagtttc tagcacatta tgacctgtt catgtgtaaa ttgggaaatg agccagaaa 2820
119 gtttattatc taattcattg attcatgcga ctatggatac atatgggatt tccagaacaa 2880
121 acagatgcaa caaagcacgg cattttccaa agatcgagtc ctcccaacag tatgcggcaa 2940
123 ggtttgttgt taagagatat aaaagcagac gacaaaacaa atcgtttatc gacctgtgc 3000
125 accaacaccg tgaacgtttg acgagtttgt agagttgtag ttgttctgt tcaaaggagc 3060
127 tccagactgg acgttccaa gcttcaacaa cttctcggca gcgtcgtgt tccgggaaa 3120
129 aaaaggcaaa aaggaacaga gcgataagca tatgtgatc tctactctt ataggctctt 3180
131 agctcaagtc aactcacatg tctttggcgg tacogaagac gttctcaagc tgcgtcttg 3240
133 aagcttttcc gagcttgcca gtaggtccct ggttgagaaa gaagatgtcg aaggctaagg 3300
135 gcgatgaaaa gcatgaagat attagctatc ggcgcgataa aagtgtgacg agatgaaaat 3360
137 ggagaaaaa tgattcgcac catcgacgac ctcgaccaa ggaatggagg tgcaccggc 3420
139 ctcccaactt ttgtaactct caacgttgac gaagatgacg aagcagtcgg tggccttagc 3480
141 ctccggctcg tagatgtga tgaaacacaa taggtatgag gagaggagaa agagaagatg 3540
143 atgagatgtc aggatgcttg cttcactgta gatggaggaa gaagatatgc gaagcaagac 3600
145 atacaactttg gaaagagctt gaaccattgt ag 3632
148 <210> SEQ ID NO: 2
149 <211> LENGTH: 3375
150 <212> TYPE: DNA
151 <213> ORGANISM: Phaffia rhodozyma
153 <400> SEQUENCE: 2
154 cttatccttc tgcgctggt ctctgtctgt cgagtgtgtg ggatggttt ggatatgttc 60
156 ctatacgaag ggtagcgcag agcaaagctg acagtattaa gcaagacaag agcttctttc 120
158 tgttgacaga tgaaaggacg aactatgaag ctgtccatgc tccccaaacc gattgacaca 180
160 ccgcgctcag gcaacgcaga atttctcact gcttcgacgt cacaccaaca tgcactctcc 240
162 atacctaaaa gcaqatcgag acacattgtt ggtcgccatg ttggatggat gtacatcaaa 300
164 cccacagcat atatcaactc catgtgagaa ctccgtagcc tctacctct tgtctctcaa 360
166 tctgaatgtc tcgttgagag gtggaatgaa tgtttacagt ttgagaagac gaaagaaaag 420
168 aagagaagag aagagaggaa tacgtacgac gaagttatca tcttatggga acttttctaa 480
170 aaaactgcct atagtagaga cgtctctctg aggaaagctc tgtagtatga tagtgaagag 540
172 cgagcaagtc tgggcaagtg cctccttctg ctacaagaaa gagaccagga aatgaaggag 600
174 agacagagtaa gcaagtaact accgatattg gatcgttctc tctaccagc gatgccttca 660

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/727,855B

DATE: 02/06/2002
 TIME: 09:42:42

Input Set : A:\PTO_VSK.txt
 Output Set: N:\CRF3\02062002\I727855B.raw

```

176 ccacgcgttc tatctcttct tgggatggca gatacatact taacgagagc aatctgatgt 726
178 ataccgaact tcgaacggaa tgatcccaga atctcttga acccttgaac ccttgaaccc 786
180 tggaaaccaag taccacccga gcaacacgcc gatacggtec acaccacaga accacacgcc 846
182 ctogtcatta aaggtgggac gcgcgatgc tggttacgtt cggcccaatc cgggaagtac 906
184 cggcttggac gtgcctgtaa ccattgcctg acggtatttc gcttcagct aactccatct 966
186 catctttttc ctttactacc acaacccacc cttgaacctt ctcccgggc ttttttacta 1026
188 tatccatcta tcaatcatca tggtcctta cactcttccc gacgtaagct taaagtttga 1086
190 gctgtgtgtg ctttatctca tcttggagtt gaactcaccg ttttttgttt ttgcttctg 1146
192 gtttttttat cggcatccct cctttttttc cctcgtggtt cgcataatgat ttgctcatca 1206
194 atcggcggtt cccatgcctc tttgtcctcc gtttcagctt ccttacgctt acgatgcctt 1266
196 ggagccttac atctctaagg gtgagattct tagtcagact gttgttccgg ttcgacacga 1326
198 tagctaatac totctcgttc ctcaatatga acatgcagaa atcatgatcc ttccaccctc 1386
200 caagcaccat cagacttacg tcttacgtta tctaaaggtc atctccgtct acatggccgg 1446
202 atcaacttgc tcatagatct tcttctgtt cggcgctacg tagaccaacc tcaacgccgc 1506
204 tatccaggct ttctcccaga ccaatgacat caaggcccag atcgtctctc agagcgtctt 1566
206 caagttcaac ggaggaggac acatcaaccg tacgatcatt ctccctcttc tggcttatca 1626
208 tatgtgttgc ttgtcactaa cagcatgca acccgggat atctcaccct gtagactccc 1686
210 tcttctggaa gaacatggtt cctgccgact ctgctgatgc caagctcacc gagggatcgc 1746
212 tcaagactgc catcgacaag gactttggat ccttcgagga gttcaagaag aagttcaaca 1806
214 ctgctactct cgggtgtccag ggatctgtca gtatctcgtt tgcttcgaca tactctcagc 1866
216 tttccttccg taaactgacg aatagttttt cggacatgta cttgtaggga tggggatggc 1926
218 tctgtcgttt gacctttttc cactttgaac attagcgata gtgataccta acaactgtga 1986
220 attggaatat agggatacaa caccgctacc aagcacctcg agatcgccac caccgccaac 2046
222 caggatcccc ttatcagtat gtgacttctc tctgttggtc accataagcc agttgctgac 2106
224 acatttcgtt cgtgtctctc cgacttcgta gctttgactc ccattcattg tcttgacgtt 2166
226 agtaattcta tctagtattt ggagtcagat tctgaacttg ccttgatctc aaacgaatga 2226
228 atcaatttct tttggtagat ctgggagcac gctttctacc tccagtacaa gaatgtcaag 2286
230 cctgattacc ttgggtacgta attctctatt cgtttgccc ggtttgatct ttgactcact 2346
232 cttaaaaatg ttttcgtttg taactttgaa aaacagccgc tttctggaac gtctgcaact 2406
234 ttgctgaggg tcagcgaagg tttgatgtga gtacaggcgc taccctacg gaggaagcga 2466
236 aggtgagctg accacttttt atctttctga tttggaatga acgatccgat gatcaaacaa 2526
238 acaggctgct gtcaaggctt aatggtccca tttatctctt tgattcgacg gcgatgacgg 2586
240 ctttctcgca tccgaagaag gcaaggctat gattactggt attctgcoat gttgcttgc 2646
242 ttgctatgct ctatgttctt ttcttttgcc tctcttcaaa gccaaaggct taaggaggc 2706
244 ccttcagctc gttttacata tgcacatata catgagaaca tatcacggac tcggcggtcg 2766
246 gtggtcctct tgagcgtcgg cttcaagatt agtgtccaca cgtgaagcgt tcgggtccat 2826
248 ccaacctggt aggaatcccc atcgggcggg aatccaatta tcaattggcg gtcggccaga 2886
250 ttcgagctcg ggtatctcag aagcgtcaag cgggcgcatt tccaggcctt taagaggagc 2946
252 aaattttaac cgccctgggtg ttcagcgaga cacgaacagt ttgaaacaga gtcgtcttgt 3006
254 gagttactcg gcgagatcac tgaggactaa actttctcag ctcgtggagc aaaagaacga 3066
256 accaaaaggc cttccctgta tctcgacctc ctcttctcc atctcttaca acactcgga 3126
258 tgaactccaa ggcttgcctt ccaaagttca aacaaactcc gggttgcat ccactgggtt 3186
260 tgtctctaac gagccgaggg atatccatcg ttgggaacgt ttgaacagac tggatggtag 3246
262 gtggccgggt gcttcggaag ccaatcataa tgggtgggaat cgagagaagg aatgattggg 3306
264 cccagtgttt aagacttgtg tttgttggca gagtacggac ggaaagtagg acagacttaa 3375
266 tcaaggcgag ccaag
269 <210> SEQ ID NO: 3
270 <211> LENGTH: 951
271 <212> TYPE: DNA

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/727,855B

DATE: 02/06/2002
TIME: 09:42:42

Input Set : A:\PTO_VSK.txt
Output Set : N:\CRF3\02062002\I727855B.raw

272 <213> ORGANISM: Phaffia rhodozyma

274 <400> SEQUENCE: 3

```

275 tccggaagct cagataccgc tcgagatcct cgagggttgt gtgctttcgc tttgttcgca      60
276 tggatgaagc tggttaactta aaaaaatcct cgtgtttctc tttgtttcaa cataggtttc      120
277 tctcttaagg tcaagacctc tgagggaaac tgggactttg taagtattct tatcgactga      180
278 gtcatacaagc tcgttatcgc tctcttaccb tcctcctttt gtgtctctgt ctacacotct      240
281 aggtcggaaga caacactccc atctttttct tgagagaccc agccaagttt ccgatcttca      300
282 ttcacaccca gaagaggaac ccgcagacaa actctaaaga caaggacgct ttctgggact      360
283 accgttcgta taaccttgct actccctgcg tgcgcctctg attcatgttg acctgtctt      420
284 tgatataatt ttatagtatc ccaaaacccc gagtcctctg atcagggtgt gcacctgttc      480
285 agtgatcgag gaacccctgc ttcttaccca cacatgcctg gttactctgg acacaccttc      540
286 aagatggtca acaggaacgg tgactggaat tatgtccaga ttcacatgcg caccgatcag      600
287 ggtgtcaaga ctcacaccaa tgaagaggct tcgaaactcg acgctcccaa tcccgaattca      660
288 aacggagacg acttggttcga cgcaatcaag aatggagact tccctagctg gacggttcag      720
289 gtacaggtaa tgtctcctga gcaggcccag aagttcagat acaacattct ggatctcacc      780
290 aaggtctggt ccacacaagga gttccactt aggaacgatt gaaagttcac ttgaaccga      840
291 aacgtggata actatttcgc agaggttgaa cagctcgctt ttgtccttc ccatctgctt      900
292 cctggaatcg agccctcgaa cgatcccgtc cttcaggctc gactattctc c
293

```

308 <210> SEQ ID NO: 4

309 <211> LENGTH: 669

310 <212> TYPE: DNA

311 <213> ORGANISM: Phaffia rhodozyma

313 <220> FEATURE:

314 <221> NAME/KEY: CDS

315 <222> LOCATION: (1)..(666)

316 <223> OTHER INFORMATION: n or X = A, C, G or T

319 <400> SEQUENCE: 4

```

320 atg tct gtt cga gca tcc ctc tct tcc gtg tct aga cag act ttc gtc      48
321 Met Ser Val Arg Ala Ser Leu Ser Ser Val Ser Arg Gln Thr Phe Val
322 1 5 10 15
323 gct cct gct gct ttc cag atc agg gca aag cat acc ctg cct gag ctt      36
324 Ala Pro Ala Ala Phe Gln Ile Arg Ala Lys His Thr Leu Pro Glu Leu
325 20 25 30
326 cct tac gct tac gat gcc ctg gag ccc tcc atc tcc aag gag atc atg      144
327 Pro Tyr Ala Tyr Asp Ala Leu Glu Pro Ser Ile Ser Lys Glu Ile Met
328 35 40 45
329 acc ctt cac cac acc aag cac cat cag act tat gtt aac ggc ctc aac      192
330 Thr Leu His His Thr Lys His His Gln Thr Tyr Val Asn Gly Leu Asn
331 50 55 60
332 gct gcc gag gag agc tac tcg gcc gct gtg ggc aag gag gat gtg ctt      240
333 Ala Ala Glu Glu Ser Tyr Ser Ala Ala Val Gly Lys Glu Asp Val Leu
334 65 70 75 80
335 acc cag gtt aag ctt cag tct gct ctc aag ttc aac gga gga gga cac      288
336 Thr Gln Val Lys Leu Gln Ser Ala Leu Lys Phe Asn Gly Gly Gly His
337 85 90 95
338 atc aat cac tct ctg ttc tgg aag aac ttg gct ccc tat gga tcc gag      336
339 Ile Asn His Ser Leu Phe Trp Lys Asn Leu Ala Pro Tyr Gly Ser Glu
340 100 105 110
341 gag gct acc ctc tct gaa gga cct ctc aag aag gct atc gag gaa tct      384
342

```

RAW SEQUENCE LISTING

DATE: 02/06/2002

PATENT APPLICATION: US/09/727,855B

TIME: 09:42:42

Input Set : A:\PTO_VSK.txt

Output Set: N:\CRF3\02062002\I727855B.raw

```

349 Glu Ala Thr Leu Ser Glu Gly Pro Leu Lys Lys Ala Ile Glu Glu Ser
350      115      120      125
352 ttt ggt tct ttc gag gcc ttc aag aag aag ttc aac ggt gac acc ggt      432
353 Phe Gly Ser Phe Glu Ala Phe Lys Lys Lys Phe Asn Ala Asp Thr Ala
354      130      135      140
356 ggt gtc caa gga tcc gga tgg gcc tgg ctt gcc ttg aac ccg ctt act      480
357 Ala Val Gln Gly Ser Gly Trp Gly Trp Leu Gly Leu Asn Pro Leu Thr
358 145      150      155      160
360 aag aag ctg gaa gtc acc acg acc gcc aac cag gac cct ctg ctt act      528
361 Lys Lys Leu Glu Val Thr Thr Thr Ala Asn Gln Asp Pro Leu Leu Thr
362      165      170      175
364 cac att cct atc atc gga gtt gac atc tgg gag cac ggt ttc tac ctt      576
365 His Ile Pro Ile Ile Gly Val Asp Ile Trp Glu His Ala Phe Tyr Leu
366      180      185      190
368 cag tac aag aac gtc aag cct gac tat ctc gct gct gtt tgg tcc gtt      624
369 Gln Tyr Lys Asn Val Lys Pro Asp Tyr Leu Ala Ala Val Trp Ser Val
370      195      200      205
372 atc aac tac aag gag gca gag gcc cga ttg cag gct gct ctc taa      669
373 Ile Asn Tyr Lys Glu Ala Glu Ala Arg Leu Gln Ala Ala Leu
374      210      215      220
377 (210)> SEQ ID NO: 5
378 (211)> LENGTH: 222
379 (212)> TYPE: PRT
380 (213)> ORGANISM: Phaffia rhodomyza
382 (400)> SEQUENCE: 5
384 Met Ser Val Arg Ala Ser Leu Ser Ser Val Ser Arg Gln Thr Phe Val
385 1      5      10      15
388 Ala Pro Ala Ala Phe Gln Ile Arg Ala Lys His Thr Leu Pro Glu Leu
389      20      25      30
392 Pro Tyr Ala Tyr Asp Ala Leu Glu Pro Ser Ile Ser Lys Glu Ile Met
393      35      40      45
396 Thr Leu His His Thr Lys His His Gln Thr Tyr Val Asn Gly Leu Asn
397      50      55      60
400 Ala Ala Glu Glu Ser Tyr Ser Ala Ala Val Gly Lys Glu Asp Val Leu
401 65      70      75      80
404 Thr Gln Val Lys Leu Gln Ser Ala Leu Lys Phe Asn Gly Gly Gly His
405      85      90      95
408 Ile Asn His Ser Leu Phe Trp Lys Asn Leu Ala Pro Tyr Gly Ser Glu
409      100      105      110
412 Glu Ala Thr Leu Ser Glu Gly Pro Leu Lys Lys Ala Ile Glu Glu Ser
413      115      120      125
416 Phe Gly Ser Phe Glu Ala Phe Lys Lys Lys Phe Asn Ala Asp Thr Ala
417      130      135      140
420 Ala Val Gln Gly Ser Gly Trp Gly Trp Leu Gly Leu Asn Pro Leu Thr
421 145      150      155      160
424 Lys Lys Leu Glu Val Thr Thr Thr Ala Asn Gln Asp Pro Leu Leu Thr
425      165      170      175
428 His Ile Pro Ile Ile Gly Val Asp Ile Trp Glu His Ala Phe Tyr Leu
429      180      185      190

```

→

VERIFICATION SUMMARY

DATE: 02/06/2002

PATENT APPLICATION: US/09/727,855B

TIME: 09:42:43

Input Set : A:\PTO_VSK.txt

Output Set: N:\CRF3\02062002\I727855B.raw

L:719 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10
L:737 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:807 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16
L:825 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17